

A hypothesis testing-based discussion on the sensitivity of means of fuzzy data with respect to data shape

María Asunción Lubiano ^{*a}, Antonia Salas^a, María Ángeles Gil^a

^a*Departamento de Estadística e I.O. y D.M.,
Universidad de Oviedo, 33071 Oviedo, Spain*

Abstract

In assessing fuzzy numbers to model imprecise data associated with random experiments, trapezoidal fuzzy numbers are often considered. Such an assessment is mainly due to easing both interpretation and computation. This becomes especially noticeable when those assessing fuzzy numbers to data have a weak knowledge, low background and little or no expertise in using fuzzy sets (as it happens when questionnaires whose responses involve a free fuzzy rating are conducted), since the required training to explain the meaning and use of trapezoidal fuzzy numbers is definitely lower than that associated with other shapes. Nevertheless, a question that constantly arises in connection with this trapezoidal assessment is whether it can importantly affect the conclusions of the study involving such data.

This paper aims to answer the last question from a statistical perspective. More concretely, the analysis of the influence of the choice of trapezoidal fuzzy numbers to model data is to be based on the conclusions from statistical hypothesis testing about the mean values of the involved fuzzy datasets. For this purpose, the p -values of tests have been compared for trapezoidal assessment *vs.* other frequently used ones, like some LU assessments. The analysis is first performed by developing simulation-based pairwise comparisons, and it is later illustrated and corroborated to some extent with a real-life example. The analysis indicates that the shape of the fuzzy assessment scarcely affects statistical conclusions.

Keywords: trapezoidal fuzzy number, LU fuzzy number, random fuzzy number, testing about means, free fuzzy rating

*corresponding author: lubiano@uniovi.es, Fax: (+34) 985103354.

1. Introduction

As recently pointed out by Grzegorzewski and Stefanini [27], fuzzy numbers have shown to play an especially relevant role in modeling and handling imprecise information, usually associated with people's cognitive states.

A problem that has been frequently investigated in connection with the use of fuzzy numbers is how to determine them. More concretely, how to assess/choose curves formalizing the imprecise valuations/perceptions, etc. As posed by Belohlavek [5]: “Do exact shapes of fuzzy sets matter?” Different responses to this question, being mostly context-dependent, can be found in the literature. In this regard, Dombi and Gera [17] pointed out that “... most fuzzy applications use piecewise linear membership functions because of their easy handling, ... In other areas where the model parameters are learned by a gradient-based optimization method, they cannot be used because of the lack of continuous derivatives.”

On one hand, in several studies related to fuzzy logical models, and according to some usual criteria in such a context, the exact shape seems not to really matter (see, for instance, Belohlavek [5]). The same applies to most of fuzzy linear programming problems for which, in accordance with different sensitivity developments (see, for instance, García-Aguado and Verdegay [19], Ebrahimnejad and Verdegay [18], and Verdegay [47]), the optimal solutions appear not to be very much affected by changes in the membership functions, and advantages associated with considering trapezoidal fuzzy numbers have been highlighted. On the other hand, many studies related to fuzzy systems (see, for instance, Šaletić *et al.* [38]) conclude that their performances are very sensitive to the shape of the membership functions. The last assertion is corroborated, among others, in the framework of fuzzy control where the suitability of using bell-shaped fuzzy numbers has been underlined by several studies (see, for instance, Baglio *et al.* [2], Koprinkova-Hristova and Penev [31], and Koprinkova [30]).

Trapezoidal (including triangular as a special case) fuzzy numbers have been proved to be a commonly welcome choice because of they showing the following skills:

- the ease to handle for most of the methodological and practical computations, whenever no differentiability of fuzzy numbers is required;
- the ease to interpret them in a natural way;

- the ease of the practical elicitation, especially when the expertise of people assessing fuzzy numbers is rather low, as it happens with children or people with no previous fuzzy set background.

In this way, a trapezoidal fuzzy number is fully determined by establishing its *core*, or 1-level set, which is the interval of the real values considered to be fully compatible with the valuation to be elicited, along with its *support* (or its closure, usually called 0-level set), which is the interval of the real values considered to be compatible to some extent with the valuation to be elicited. The remaining values can be directly obtained by linearly interpolating from these two intervals.

Therefore, the intuition behind the trapezoidal representation is rather easily caught. This fact becomes especially useful when people in charge of fuzzy assessments are not familiarized enough with the concepts around, like it often happens, for instance, in dealing with questionnaires about opinion/satisfaction/judgement/... (see, e.g., Hesketh *et al.* [29], Hesketh and Hesketh [28], Colubi *et al.* [11], Denoeux [15], González-Rodríguez *et al.* [22], Gil *et al.* [20], and Lubiano *et al.* [34]).

It should be also pointed out that a topic receiving a wide attention in recent years is the approximation (w.r.t. certain metrics) of fuzzy numbers by means of trapezoidal ones, often preserving some key characteristics, like the expected interval, the ambiguity, and so on (see, among others, Abbasbandy and Asady [1], Grzegorzewski and Mrówka [25], Grzegorzewski [24], Yeh [49, 50], Ban and Coroianu [3, 4], Grzegorzewski and Pasternak-Winiarska [26] and Coroianu and Stefanini [12]).

In this paper, we are going to show empirically that when fuzzy numbers are modelling imprecise data coming from random experiments and fuzzy datasets are summarized by their means, data shape is usually not relevant. The discussion is carried out on the basis of a methodology for testing hypothesis about means with fuzzy data and it is accomplished in two ways, both of them assuming fuzzy data moving in a free fuzzy scale.

Firstly, and after recalling some preliminaries in Section 2, two simulation studies are developed in Section 3. These studies concern the computation of the two-sample test about means p -values (Montenegro *et al.* [35]) when one of the two samples refers to trapezoidal fuzzy data and the other one refers to data from another class of LR fuzzy numbers (actually, LU fuzzy numbers chosen from the families in Stefanini *et al.* [45], Sorini and Stefanini [42], and Stefanini and Sorini [44]). Simulations have been accomplished according to

two different models. Although realistic, successful and popular models for the distributions of the random mechanisms generating fuzzy data cannot be found in the literature yet, we are going to consider two different ones that have been used for purposes of statistical comparisons in recent studies (see, for instance, the one by Sinova *et al.* [39, 40, 41] and the one in De la Rosa de Saa *et al.* [13], this last one mimicking to a large extent the assessments in different real-life examples involving the fuzzy rating scale by Hesketh *et al.* [29]).

Secondly, and once it is empirically shown that differences in shape of fuzzy data do not generally entail significant differences, the real-life example analyzed in detail in Gil *et al.* [20] and Lubiano *et al.* [34] is considered to perform an additional discussion in Section 4. Some of the one-, two- and k -samples hypothesis test about means p -values in Lubiano *et al.* [34] have been computed by considering all data as either being trapezoidal (as assumed in [34]), or belonging to any of the classes involved in the discussion in Section 3. Conclusions agree with those in Section 3.

The paper ends with some comments about some practical implications and open problems.

2. Preliminaries

Fuzzy numbers (also referred to as compact fuzzy intervals) will be intended along the paper as follows:

Definition 2.1. A **fuzzy number** is a mapping $\tilde{U} : \mathbb{R} \rightarrow [0, 1]$ such that for each $\alpha \in [0, 1]$ its α -level set

$$\tilde{U}_\alpha = \begin{cases} \{x \in \mathbb{R} : \tilde{U}(x) \geq \alpha\} & \text{if } \alpha \in (0, 1] \\ \text{cl}\{x \in \mathbb{R} : \tilde{U}(x) > 0\} & \text{if } \alpha = 0 \end{cases}$$

(with cl = topological closure) is a nonempty compact interval. The space of fuzzy numbers will be denoted by $\mathcal{F}_c(\mathbb{R})$.

The *core* of \tilde{U} , or 1-level set, \tilde{U}_1 , corresponds to the interval (in particular it can reduce to a singleton) of real numbers which are considered to be ‘fully compatible’ with (the property associated with) \tilde{U} . The *support* of \tilde{U} (its closure being the 0-level set, \tilde{U}_0), corresponds to the interval of real numbers which are considered to be ‘compatible to some extent’ with \tilde{U} .

In developing statistics with fuzzy data, the two following extended operations based on Zadeh’s extension principle [51] should be considered:

Definition 2.2. If $\tilde{U}, \tilde{V} \in \mathcal{F}_c(\mathbb{R})$, then the **sum of \tilde{U} and \tilde{V}** is defined as the fuzzy number $\tilde{U} + \tilde{V} \in \mathcal{F}_c(\mathbb{R})$ such that for each $\alpha \in [0, 1]$

$$(\tilde{U} + \tilde{V})_\alpha = [\inf \tilde{U}_\alpha + \inf \tilde{V}_\alpha, \sup \tilde{U}_\alpha + \sup \tilde{V}_\alpha].$$

If $\tilde{U} \in \mathcal{F}_c(\mathbb{R})$ and $\gamma \in \mathbb{R}$, the **product of \tilde{U} by the scalar γ** is defined as the fuzzy number $\gamma \cdot \tilde{U} \in \mathcal{F}_c(\mathbb{R})$ such that for each $\alpha \in [0, 1]$

$$(\gamma \cdot \tilde{U})_\alpha = \gamma \cdot \tilde{U}_\alpha = \{\gamma \cdot y : y \in \tilde{U}_\alpha\} = \begin{cases} [\gamma \inf \tilde{U}_\alpha, \gamma \sup \tilde{U}_\alpha] & \text{if } \gamma \geq 0, \\ [\gamma \sup \tilde{U}_\alpha, \gamma \inf \tilde{U}_\alpha] & \text{if } \gamma < 0. \end{cases}$$

It is well-known that when $\mathcal{F}_c(\mathbb{R})$ is endowed with the preceding operations, the space of fuzzy numbers is semilinear. In fact, it is not possible to establish a difference operator between fuzzy numbers that is always well-defined and preserves all the properties of the difference of real numbers. In testing hypothesis about the equality of fuzzy-valued parameters, the use of differences between their sample counterparts in the real-valued case has been replaced (see, for instance, Blanco *et al.* [7, 8] and Lubiano *et al.* [34]) by that of convenient metrics between fuzzy numbers, like the L^2 one by Diamond and Kloeden [16] which levelwise extends Vitale's one [48] for compact sets.

Definition 2.3. Let $\tilde{U}, \tilde{V} \in \mathcal{F}_c(\mathbb{R})$. The **L^2 Diamond and Kloeden distance** between \tilde{U} and \tilde{V} is given by

$$\rho_2(\tilde{U}, \tilde{V}) = \sqrt{\frac{1}{2} \int_{[0,1]} \left([\inf \tilde{U}_\alpha - \inf \tilde{V}_\alpha]^2 + [\sup \tilde{U}_\alpha - \sup \tilde{V}_\alpha]^2 \right) d\alpha}.$$

When fuzzy data are analyzed from an inferential perspective, the random mechanism generating these data associated with a random experiment can be properly formalized within the probabilistic setting by means of the so-called fuzzy random variables by Puri and Ralescu [37] (see also Kruse and Meyer [33]), that we will refer to as random fuzzy numbers in case of taking values on $\mathcal{F}_c(\mathbb{R})$.

Consider a random experiment which is mathematically modeled by means of a probability space (Ω, \mathcal{A}, P) .

Definition 2.4. A **random fuzzy number** (for short **RFN**) associated with (Ω, \mathcal{A}, P) is a mapping $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c(\mathbb{R})$ such that for all $\alpha \in [0, 1]$ the α -level mapping \mathcal{X}_α is a compact random interval; that is, for all $\alpha \in [0, 1]$ the real-valued mappings $\inf \mathcal{X}_\alpha$ and $\sup \mathcal{X}_\alpha$ (with $\mathcal{X}_\alpha(\omega) = (\mathcal{X}(\omega))_\alpha$) are random variables.

It is known (see, for instance, Colubi *et al.* [10]) that a mapping $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c(\mathbb{R})$ is an RFN if and only if it is Borel-measurable w.r.t. the σ -field generated by the topology induced by ρ_2 on $\mathcal{F}_c(\mathbb{R})$. This equivalent formalization allows us to rigorously refer to the (induced from P) distribution of the RFN \mathcal{X} , the independence of several RFNs, and so on. It also allows us to extend or preserve almost all the concepts and ideas from statistical inference with real-valued data. Consequently, with the use of random fuzzy sets one can properly refer to their (induced) parameters, the p -value and consistency of a test, the unbiasedness and consistency of an estimator, and so on.

In performing inferential analysis about the distribution of RFNs, the best known involved parameter is the Aumann-type mean value by Puri and Ralescu [37], which is defined as follows:

Definition 2.5. *Let (Ω, \mathcal{A}, P) be the probability space modeling a random experiment and \mathcal{X} be an associated RFN. The **Aumann-type mean of \mathcal{X}** is the fuzzy number $\tilde{E}(\mathcal{X}) \in \mathcal{F}_c(\mathbb{R})$, if it exists, such that for all $\alpha \in [0, 1]$*

$$\left(\tilde{E}(\mathcal{X})\right)_\alpha = [E(\inf \mathcal{X}_\alpha), E(\sup \mathcal{X}_\alpha)],$$

with E denoting the mean value of the corresponding real-valued random variable.

To test the null hypothesis that the Aumann-type mean of an RFN \mathcal{X} equals a given fuzzy number \tilde{U} , one can consider the **bootstrapped algorithm to one-sample test about the mean of an RFN** by González-Rodríguez *et al.* [23] (see also Colubi [9]), approximating the asymptotic approaches by Körner [32] and Montenegro *et al.* [36], which is now algorithmically summarized. Let $\tilde{\mathbf{x}}_n = (\tilde{x}_1, \dots, \tilde{x}_n)$ be a sample of independent observations from the RFN \mathcal{X} , and let $\tilde{U} \in \mathcal{F}_c(\mathbb{R})$. Then, to test the null hypothesis $H_0 : \tilde{E}(\mathcal{X}) = \tilde{U}$ we can proceed following the one-sample algorithm:

Step 1. Compute the value of the statistic

$$T_n = \frac{A_n(\tilde{\mathbf{x}}_n, \tilde{U})}{B_n(\tilde{\mathbf{x}}_n)} = \frac{\left[\rho_2\left(\frac{1}{n} \cdot (\tilde{x}_1 + \dots + \tilde{x}_n), \tilde{U}\right)\right]^2}{\frac{1}{n-1} \sum_{i=1}^n \left[\rho_2\left(\tilde{x}_i, \frac{1}{n} \cdot (\tilde{x}_1 + \dots + \tilde{x}_n)\right)\right]^2} \quad (1)$$

Step 2. Fix the bootstrap population to be $\{\tilde{x}_1, \dots, \tilde{x}_n\}$.

Step 3. Obtain a sample of independent observations from the bootstrap population, $\tilde{\mathbf{x}}_n^* = (\tilde{x}_1^*, \dots, \tilde{x}_n^*)$.

Step 4. Compute the value of the bootstrap statistic

$$T_n^* = \frac{A_n\left(\tilde{\mathbf{x}}_n^*, \frac{1}{n} \cdot (\tilde{x}_1 + \dots + \tilde{x}_n)\right)}{B_n(\tilde{\mathbf{x}}_n^*)} \quad (2)$$

Step 5. Steps 3 and 4 should be repeated a large number W of times to get a set of estimates, denoted by $\{t_1^*, \dots, t_W^*\}$.

Step 6. Compute the bootstrap p -value as the proportion of values in $\{t_1^*, \dots, t_W^*\}$ which are greater than T_n .

To test the null hypothesis of equality of the Aumann-type means of two independent RFNs, \mathcal{X} and \mathcal{Y} , one can consider the ***bootstrapped algorithm to two-sample test about the equality of means of two independent RFNs*** in Colubi [9], approximating the procedure by Montenegro *et al.* [35], which is now algorithmically summarized. If \mathcal{X} and \mathcal{Y} are two independent RFNs, consider a sample of independent observations from \mathcal{X} , $\tilde{\mathbf{x}}_{n_1} = (\tilde{x}_1, \dots, \tilde{x}_{n_1})$, and a sample of independent observations from \mathcal{Y} , $\tilde{\mathbf{y}}_{n_2} = (\tilde{y}_1, \dots, \tilde{y}_{n_2})$. The algorithm to test $H_0 : \tilde{E}(\mathcal{X}) = \tilde{E}(\mathcal{Y})$ proceeds as follows:

Step 1. Compute the value of the statistic

$$T_{n_1, n_2} = \frac{A_{n_1, n_2}(\tilde{\mathbf{x}}_{n_1}, \tilde{\mathbf{y}}_{n_2})}{\frac{B_{n_1}(\tilde{\mathbf{x}}_{n_1})}{n_1} + \frac{B_{n_2}(\tilde{\mathbf{y}}_{n_2})}{n_2}} \quad (3)$$

where

$$A_{n_1, n_2}(\tilde{\mathbf{x}}_{n_1}, \tilde{\mathbf{y}}_{n_2}) = \left[\rho_2\left(\frac{1}{n_1} \cdot (\tilde{x}_1 + \dots + \tilde{x}_{n_1}), \frac{1}{n_2} \cdot (\tilde{y}_1 + \dots + \tilde{y}_{n_2})\right) \right]^2,$$

$$B_{n_1}(\tilde{\mathbf{x}}_{n_1}) = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} \left[\rho_2\left(\tilde{x}_i, \frac{1}{n_1} \cdot (\tilde{x}_1 + \dots + \tilde{x}_{n_1})\right) \right]^2,$$

$$B_{n_2}(\tilde{\mathbf{y}}_{n_2}) = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} \left[\rho_2\left(\tilde{y}_j, \frac{1}{n_2} \cdot (\tilde{y}_1 + \dots + \tilde{y}_{n_2})\right) \right]^2.$$

Step 2. Fix the bootstrap populations to be as follows (to ensure that bootstrap populations fulfill the null hypothesis, one can add to each value in each sample the mean of the other one):

$$\left\{ \tilde{x}_1 + \frac{1}{n_2} \cdot (\tilde{y}_1 + \dots + \tilde{y}_{n_2}), \dots, \tilde{x}_{n_1} + \frac{1}{n_2} \cdot (\tilde{y}_1 + \dots + \tilde{y}_{n_2}) \right\},$$

$$\left\{ \tilde{y}_1 + \frac{1}{n_1} \cdot (\tilde{x}_1 + \dots + \tilde{x}_{n_1}), \dots, \tilde{y}_{n_2} + \frac{1}{n_1} \cdot (\tilde{x}_1 + \dots + \tilde{x}_{n_1}) \right\}.$$

Step 3. Obtain a sample of independent observations from each bootstrap population, $\tilde{\mathbf{x}}_{n_1}^* = (\tilde{x}_1^*, \dots, \tilde{x}_{n_1}^*)$ and $\tilde{\mathbf{y}}_{n_2}^* = (\tilde{y}_1^*, \dots, \tilde{y}_{n_2}^*)$.

Step 4. Compute the value of the bootstrap statistic

$$T_{n_1, n_2}^* = \frac{A_{n_1, n_2}(\tilde{\mathbf{x}}_{n_1}^*, \tilde{\mathbf{y}}_{n_2}^*)}{\frac{B_{n_1}(\tilde{\mathbf{x}}_{n_1}^*)}{n_1} + \frac{B_{n_2}(\tilde{\mathbf{y}}_{n_2}^*)}{n_2}} \quad (4)$$

Step 5. Steps 3 and 4 should be repeated a large number W of times to get a set of estimates, denoted by $\{t_1^*, \dots, t_W^*\}$.

Step 6. Compute the bootstrap p -value as the proportion of values in $\{t_1^*, \dots, t_W^*\}$ being greater than T_{n_1, n_2} .

To test the null hypothesis of equality of the Aumann-type means of k independent RFNs, $\mathcal{X}_1, \dots, \mathcal{X}_k$, one can consider the ***bootstrapped algorithm to oneway ANOVA test about the equality of means of k independent RFNs*** in González-Rodríguez *et al.* [22], extending the procedure by Gil *et al.* [21], which is now algorithmically summarized. If $\mathcal{X}_1, \dots, \mathcal{X}_k$ are independent RFNs, consider a sample of independent observations $\tilde{\mathbf{x}}_j = (\tilde{x}_{1j}, \dots, \tilde{x}_{n_j j})$ from \mathcal{X}_j , $j = 1, \dots, k$, the k samples being also independent. Denote $\tilde{\mathbf{x}}_{n_1 + \dots + n_k} \equiv (\tilde{x}_{11}, \dots, \tilde{x}_{n_1 1}, \dots, \tilde{x}_{1k}, \dots, \tilde{x}_{n_k k})$, $\tilde{\mathbf{x}}_j = \frac{1}{n_j} \cdot (\tilde{x}_{1j} + \dots + \tilde{x}_{n_j j})$ and $\tilde{\mathbf{x}}_{(-j)} = \tilde{\mathbf{x}}_1 + \dots + \tilde{\mathbf{x}}_{j-1} + \tilde{\mathbf{x}}_{j+1} + \dots + \tilde{\mathbf{x}}_k$. Then, the algorithm to test the null hypothesis $H_0 : \tilde{E}(\mathcal{X}_1) = \dots = \tilde{E}(\mathcal{X}_k)$ proceeds as follows:

Step 1. Compute the value of the statistic

$$T_{n_1, \dots, n_k} = \frac{\sum_{j=1}^k n_j \cdot A_{n_j, n_1 + \dots + n_k}(\tilde{\mathbf{x}}_j, \tilde{\mathbf{x}}_{n_1 + \dots + n_k})}{\sum_{j=1}^k \frac{n_j - 1}{n_j} B_{n_j}(\tilde{\mathbf{x}}_j)}$$

$$\begin{aligned}
& \sum_{j=1}^k n_j \left[\rho_2 \left(\frac{1}{n_j} \cdot (\tilde{x}_{1j} + \dots + \tilde{x}_{n_j j}), \frac{1}{n_1 + \dots + n_k} \cdot (\tilde{x}_{11} + \dots + \tilde{x}_{n_k k}) \right) \right]^2 \\
= & \frac{\sum_{j=1}^k \frac{1}{n_j} \sum_{i=1}^{n_j} \left[\rho_2 \left(\tilde{x}_{ij}, \frac{1}{n_j} \cdot (\tilde{x}_{1j} + \dots + \tilde{x}_{n_j j}) \right) \right]^2}{\phantom{\sum_{j=1}^k \frac{1}{n_j} \sum_{i=1}^{n_j} \left[\rho_2 \left(\tilde{x}_{ij}, \frac{1}{n_j} \cdot (\tilde{x}_{1j} + \dots + \tilde{x}_{n_j j}) \right) \right]^2}} \quad (5)
\end{aligned}$$

Step 2. Fix the bootstrap populations to be $\left\{ \tilde{x}_{1j} + \overline{\tilde{\mathbf{x}}_{(-j)}}, \dots, \tilde{x}_{n_j j} + \overline{\tilde{\mathbf{x}}_{(-j)}} \right\}$, for each $j = 1, \dots, k$ (to get bootstrap populations fulfilling the null hypothesis).

Step 3. Obtain a sample of independent observations from each bootstrap population, $\tilde{\mathbf{x}}_j^* = (\tilde{x}_{1j}^*, \dots, \tilde{x}_{n_j j}^*)$, $j = 1, \dots, k$, and denote $\tilde{\mathbf{x}}_{n_1 + \dots + n_k}^* = (\tilde{x}_{11}^*, \dots, \tilde{x}_{n_1 1}^*, \dots, \tilde{x}_{1k}^*, \dots, \tilde{x}_{n_k k}^*)$.

Step 4. Compute the value of the bootstrap statistic

$$T_{n_1, \dots, n_k}^* = \frac{\sum_{j=1}^k n_j \cdot A_{n_j, n_1 + \dots + n_k}(\tilde{\mathbf{x}}_j^*, \tilde{\mathbf{x}}_{n_1 + \dots + n_k}^*)}{\sum_{j=1}^k \frac{n_j - 1}{n_j} B_{n_j}(\tilde{\mathbf{x}}_j^*)} \quad (6)$$

Step 5. Steps 3 and 4 should be repeated a large number W of times to get a set of estimates, denoted by $\{t_1^*, \dots, t_W^*\}$.

Step 6. Compute the bootstrap p -value as the proportion of values in $\{t_1^*, \dots, t_W^*\}$ being greater than T_{n_1, \dots, n_k}^* .

The preceding algorithms have been implemented in the R package SAFD (Statistical Analysis of Fuzzy Data, see Trutschnig and Lubiano [46]). In Lubiano *et al.* [34], we can find the particularization of the preceding algorithms when all the involved data are trapezoidal.

3. Hypothesis testing-based empirical analysis of the sensitivity of means of RFNs w.r.t. changes in shape

This section deals with the analysis of the influence of the choice of the shape of the membership functions of fuzzy data on the Aumann-type mean of such data. More concretely, this analysis is to be focussed on discussing whether choosing trapezoidal fuzzy numbers to model fuzzy data associated with random experiments leads to significantly different mean values that making some other usual choices.

Certainly, the lack of realistic models for the distributions of random fuzzy numbers prevents us to establish general conclusions. Nevertheless, rather broad conclusions can be drawn by developing simulation studies and applying pairwise comparisons of trapezoidal *vs* the other considered types of data. To this purpose, the pairwise comparison is to be based on the bootstrapped two-sample test about the equality of means in Section 2.

The types of fuzzy numbers to be involved in this section correspond to some of the LU s (see, for instance, Stefanini *et al.* [45], Sorini and Stefanini [42], Stefanini and Bede [43]) that can be characterized by means of four real numbers (namely, the extremes of their core and support).

In particular, we have chosen trapezoidal fuzzy numbers (Tra) *vs* quadratic functions (Π -curves), functions with parametric monotonic Hermite-type interpolation either using (2,2)-rational splines (LU_{1A} and LU_{1B}), or mixed exponential splines (LU_{2A} and LU_{2B}) (see Figure 1). More specifically, if $\tilde{U} \equiv LU(a, b, c, d)$ with $a = \inf \tilde{U}_0$, $b = \inf \tilde{U}_1$, $c = \sup \tilde{U}_1$, $d = \sup \tilde{U}_0$, and $LU \in \{\Pi, LU_{1A}, LU_{1B}, LU_{2A}, LU_{2B}\}$, then for each $\alpha \in [0, 1]$

$$\tilde{U}_\alpha = [a + l_{LU}(\alpha)(b - a), c + r_{LU}(\alpha)(d - c)],$$

where the functions involved in the left and right arms can be found in Table 1.

Table 1: Expressions for functions l_{LU} and r_{LU} in the horizontal view of LU fuzzy numbers with LU ranging on $\{\text{Tra}, \Pi, LU_{1A}, LU_{1B}, LU_{2A}, LU_{2B}\}$

LU	$l_{LU}(\alpha)$	$r_{LU}(\alpha)$
Tra	α	$1 - \alpha$
Π	$\begin{cases} \sqrt{\alpha/2} & \text{if } \alpha < 1/2 \\ 1 - \sqrt{(1-\alpha)/2} & \text{otherwise} \end{cases}$	$\begin{cases} 1 - \sqrt{\alpha/2} & \text{if } \alpha < 1/2 \\ \sqrt{(1-\alpha)/2} & \text{otherwise} \end{cases}$
LU_{1A}	$\frac{\alpha^2 + 5\alpha(1-\alpha)}{1 + 3.5\alpha(1-\alpha)}$	$(1-\alpha)(1+0.9\alpha)$
LU_{1B}	α	$1 - \frac{\alpha^2 + 5\alpha(1-\alpha)}{1 + 3.2\alpha(1-\alpha)}$
LU_{2A}	$\frac{\alpha^2(3-2\alpha) - 0.5(1-\alpha)^{1.55} + 0.5 + 0.05\alpha^{1.55}}{1.55}$	$1 - \frac{\alpha^2(3-2\alpha) - 5(1-\alpha)^{11} + 5 + 5\alpha^{11}}{11}$
LU_{2B}	$\frac{\alpha^2(3-2\alpha) - 0.5(1-\alpha)^{1.55} + 0.5 + 0.05\alpha^{1.55}}{1.55}$	$1 - \frac{\alpha^2(3-2\alpha) - 5(1-\alpha)^{6.05} + 5 + 0.05\alpha^{6.05}}{6.05}$

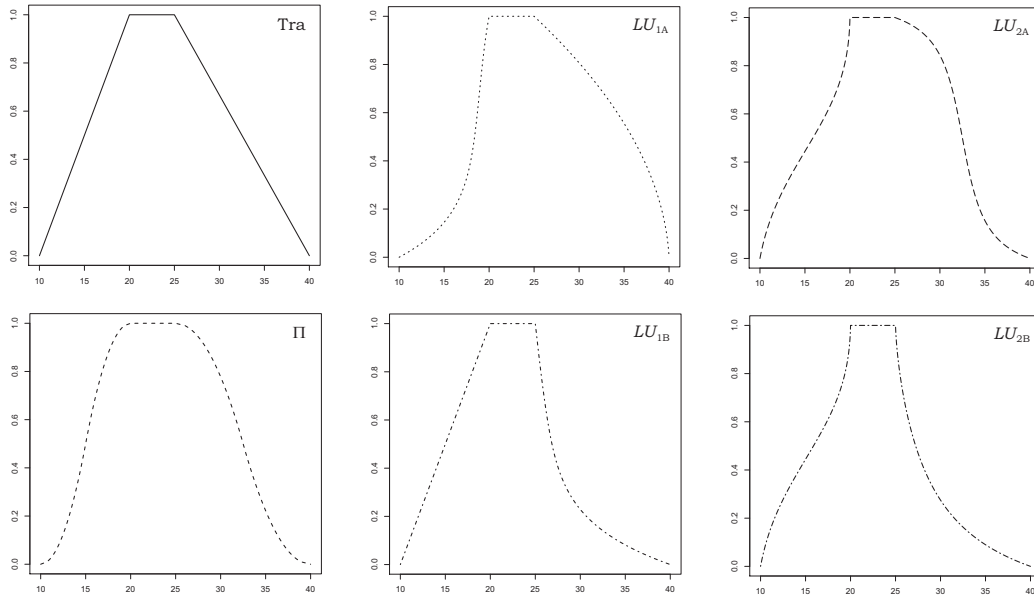


Figure 1: Six types of fuzzy numbers sharing core $[20, 25]$ and support $(10, 40)$ and differing in shape. On the left, trapezoidal (top) and Π -curve (bottom), along with four different LU fuzzy numbers on the middle and the right

Two additional fuzzy numbers have been added to the comparison, namely (see Figure 2), triangular (Tri),

$$\text{Tri}(a, b, c, d) = \text{Tra}(a, (b + c)/2, (b + c)/2, d),$$

and symmetric triangular (TriS),

$$\text{TriS}(a, b, c, d) = \text{Tra}(a, (a + d)/2, (a + d)/2, d).$$

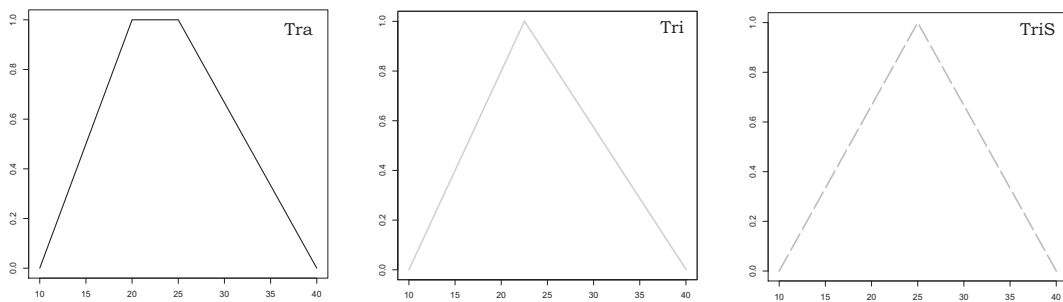


Figure 2: Trapezoidal (on the left) along with triangular (on the middle) and symmetric triangular (on the right) fuzzy numbers sharing support $(10, 40)$ although different in core.

Two key common guidelines in the simulation studies are the following:

- to generate fuzzy data, we generate values from four real-valued random variables X_1, X_2, X_3 and X_4 so that the $\mathbb{R} \times [0, \infty) \times [0, \infty) \times [0, \infty)$ -valued random vector (X_1, X_2, X_3, X_4) will provide us with the 4-tuples $(x_1, x_2, x_3, x_4) = ((b+c)/2, (c-b)/2, b-a, d-c)$ (i.e., as they are usually known, $x_1/x_2 = \text{center}/\text{radius}$ of the core, and $x_3/x_4 = \text{lower}/\text{upper}$ spread of the fuzzy number, see Figure 3);

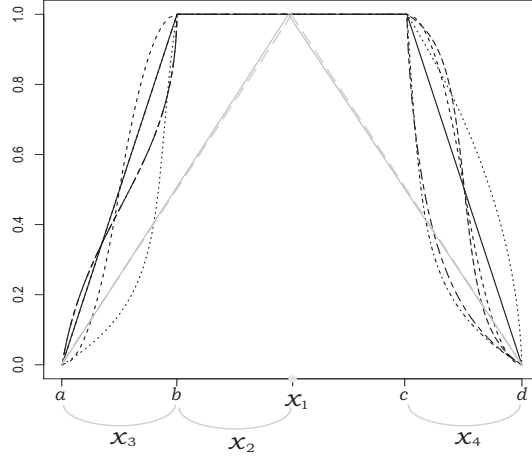


Figure 3: Different LU -fuzzy numbers associated with the same 4-tuples (a, b, c, d) and (x_1, x_2, x_3, x_4)

- to each generated 4-tuple (x_1, x_2, x_3, x_4) we associate the eight fuzzy numbers indistinctly denoted as

$$LU\langle x_1, x_2, x_3, x_4 \rangle = LU(x_1 - x_2 - x_3, x_1 - x_2, x_1 + x_2, x_1 + x_2 + x_4)$$

with $LU \in \{\text{Tra}, \Pi, LU_{1A}, LU_{1B}, LU_{2A}, LU_{2B}, \text{Tri}, \text{TriS}\}$.

On the other hand, one should take into account that

- whenever $LU \in \{\text{Tra}, \Pi, LU_{1A}, LU_{1B}, LU_{2A}, LU_{2B}, \text{Tri}, \text{TriS}\}$, we have that (see Stefanini *et al.* [45])

$$\begin{aligned} & \frac{1}{n} \cdot [LU(a_1, b_1, c_1, d_1) + \cdots + LU(a_n, b_n, c_n, d_n)] \\ &= LU\left(\frac{1}{n} \sum_{i=1}^n a_i, \frac{1}{n} \sum_{i=1}^n b_i, \frac{1}{n} \sum_{i=1}^n c_i, \frac{1}{n} \sum_{i=1}^n d_i\right); \end{aligned}$$

- in computing the values of the statistic $T_{n,n}$ in *Step 1* of the bootstrapped algorithm to two-sample test in Section 2 when it is applied to compare means of trapezoidal and LU fuzzy data (with $LU \in \{\Pi, LU_{1A}, LU_{1B}, LU_{2A}, LU_{2B}, \text{Tri}, \text{TriS}\}$), one should make use of
 - distances between $\text{Tra}(a, b, c, d)$ and $LU(a, b, c, d)$ for the numerator, whose exact expressions are gathered in Table 2;

Table 2: Exact expressions for the ρ_2 distances between fuzzy numbers $\text{Tra}(a, b, c, d)$ and $LU(a, b, c, d)$ with LU ranging on $\{\Pi, LU_{1A}, LU_{1B}, LU_{2A}, LU_{2B}, \text{Tri}, \text{TriS}\}$

LU	$[\rho_2(\text{Tra}(a, b, c, d), LU(a, b, c, d))]^2$
Π	$0.00416667(a-b)^2 + 0.00416667(c-d)^2$
LU_{1A}	$0.03129046(a-b)^2 + 0.0135(c-d)^2$
LU_{1B}	$0.03724265(c-d)^2$
LU_{2A}	$0.00171940(a-b)^2 + 0.01325610(c-d)^2$
LU_{2B}	$0.00171940(a-b)^2 + 0.05307135(c-d)^2$
Tri	$(c-b)^2/12$
TriS	$[(b-a+c-d)^2 + (c-b)^2]/12$

- distances between $LU(a_1, b_1, c_1, d_1)$ and $LU(a_2, b_2, c_2, d_2)$ involved in the denominator, where $LU \in \{\text{Tra}, \Pi, LU_{1A}, LU_{1B}, LU_{2A}, LU_{2B}\}$, whose exact expressions are gathered in Table 3;

Table 3: Exact expressions for the ρ_2 distances between fuzzy numbers $LU(a_1, b_1, c_1, d_1)$ and $LU(a_2, b_2, c_2, d_2)$ with LU ranging on $\{\text{Tra}, \Pi, LU_{1A}, LU_{1B}, LU_{2A}, LU_{2B}\}$

LU	$[\rho_2(LU(a_1, b_1, c_1, d_1), LU(a_2, b_2, c_2, d_2))]^2$
Tra	$[(a_1 - a_2)^2 + (b_1 - b_2)^2 + (a_1 - a_2)(b_1 - b_2)]/6$ $+[(c_1 - c_2)^2 + (d_1 - d_2)^2 + (c_1 - c_2)(d_1 - d_2)]/6$
Π	$[7(a_1 - a_2)^2 + 7(b_1 - b_2)^2 + 10(a_1 - a_2)(b_1 - b_2)]/48$ $+ [7(c_1 - c_2)^2 + 7(d_1 - d_2)^2 + 10(c_1 - c_2)(d_1 - d_2)]/48$
LU_{1A}	$0.06622634(a_1 - a_2)^2 + 0.2900244(b_1 - b_2)^2 + 0.14374922(a_1 - a_2)(b_1 - b_2)$ $+ 0.10516667(c_1 - c_2)^2 + 0.25516667(d_1 - d_2)^2 + 0.13966667(c_1 - c_2)(d_1 - d_2)$
LU_{1B}	$0.16666666(a_1 - a_2)^2 + 0.16666666(b_1 - b_2)^2 + 0.16666666(a_1 - a_2)(b_1 - b_2)$ $+ 0.30881332(c_1 - c_2)^2 + 0.06176266(d_1 - d_2)^2 + 0.12942402(c_1 - c_2)(d_1 - d_2)$
LU_{2A}	$0.16429256(a_1 - a_2)^2 + 0.19560186(b_1 - b_2)^2 + 0.14010559(a_1 - a_2)(b_1 - b_2)$ $+ 0.13773163(c_1 - c_2)^2 + 0.13773163(d_1 - d_2)^2 + 0.22453675(c_1 - c_2)(d_1 - d_2)$
LU_{2B}	$0.16429256(a_1 - a_2)^2 + 0.19560186(b_1 - b_2)^2 + 0.14010559(a_1 - a_2)(b_1 - b_2)$ $+ 0.34394341(c_1 - c_2)^2 + 0.05090666(d_1 - d_2)^2 + 0.10514993(c_1 - c_2)(d_1 - d_2)$

by using these exact expressions for distances, the value of the statistic $T_{n,n}$ will be computed by adapting the algorithm implemented in the SAFD package [46];

- in computing the values of the bootstrapped statistic $T_{n,n}^*$ in *Step 4* of the two-sample test in Section 2 when it is applied to compare means of trapezoidal and the other LU fuzzy data, the involved bootstrap fuzzy data will not be purely LU ones, but the output of adding a trapezoidal and another LU fuzzy number; consequently, exact expressions for the distances are cumbersome, and we will run the SAFD package [46].

The common *scheme of the multiple comparison analyses* adapting the two-sample algorithm can be summarized as follows:

S.1. A sample of n 4-tuples $\{(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, x_4^{(i)})\}_{i=1}^n$ is generated by the selected simulation procedure. To each of the 4-tuples we can associate eight samples of size n , $\{LU\langle x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, x_4^{(i)} \rangle\}_{i=1}^n$, of fuzzy data, where $LU \in \{\text{Tra}, \Pi, LU_{1A}, LU_{1B}, LU_{2A}, LU_{2B}, \text{Tri}, \text{TriS}\}$. Then, to test the null hypothesis of the equality of the means of the random fuzzy numbers $\text{Tra}\langle X_1, X_2, X_3, X_4 \rangle$ and $LU\langle X_1, X_2, X_3, X_4 \rangle$ (that is, $H_0 : \tilde{E}(\text{Tra}\langle X_1, X_2, X_3, X_4 \rangle) = \tilde{E}(LU\langle X_1, X_2, X_3, X_4 \rangle)$) for each of $LU \in \{\Pi, LU_{1A}, LU_{1B}, LU_{2A}, LU_{2B}, \text{Tri}, \text{TriS}\}$, the values of the seven statistics $T_{n,n}$ (as in equation (3)) are computed.

S.2. For each of the seven LU types different from Tra, the bootstrap populations would be

$$\left\{ \text{Tra} \langle x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, x_4^{(i)} \rangle + LU \left\langle \sum_{j=1}^n \frac{x_1^{(j)}}{n}, \sum_{j=1}^n \frac{x_2^{(j)}}{n}, \sum_{j=1}^n \frac{x_3^{(j)}}{n}, \sum_{j=1}^n \frac{x_4^{(j)}}{n} \right\rangle \right\}_{i=1}^n,$$

$$\left\{ LU \langle x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, x_4^{(i)} \rangle + \text{Tra} \left\langle \sum_{j=1}^n \frac{x_1^{(j)}}{n}, \sum_{j=1}^n \frac{x_2^{(j)}}{n}, \sum_{j=1}^n \frac{x_3^{(j)}}{n}, \sum_{j=1}^n \frac{x_4^{(j)}}{n} \right\rangle \right\}_{i=1}^n.$$

S.3. Obtain a sample of independent observations from each bootstrap population.

S.4. Compute the value of the bootstrap statistic $T_{n,n}^*$ (as in Equation (4)).

S.5. Repeat 1000 times *S.3* to get a set of statistic values, denoted by $\{t_1^*, \dots, t_{1000}^*\}$.

S.6. Compute the bootstrap p -value as the proportion of values in $\{t_1^*, \dots, t_{1000}^*\}$ being greater than the value of $T_{n,n}$.

Remark 3.1. It should be pointed out that, although the sample of Tra's and the sample of LU 's in **S.1** share the same sample of 4-tuples to avoid differences in mean coming from differences in the starting 4-tuples, they should not be treated as linked since they are actually assumed to come from two independent samples of people each sample being associated with one of the two different shapes. Of course, samples in **S.3** do not share in general the same sample of 4-tuples.

As we have announced before, the *simulation developments* have been carried out in two ways.

According to the FIRST SIMULATION PROCEDURE, data have been generated from random fuzzy numbers with an unbounded reference set and having a kind of 'normal' distribution. More concretely, from random vectors (X_1, X_2, X_3, X_4) for which X_1 is normally distributed, similar to some ones considered by Sinova *et al.* (see [39], [40]), so that $X_1 \sim \mathcal{N}(0, 1)$ and $X_2, X_3, X_4 \sim \chi_1^2$, all of them being independent.

The scheme **S.1** to **S.6** has been run in both simulation studies for 30 samples of size $n = 100$. Results from this FIRST PROCEDURE are shown in Table 4.

The p -values in testing the equality of means of trapezoidal data *vs* either Π , LU_{1A} , LU_{1B} , LU_{2A} , and LU_{2B} are high in this table. However, they are

Table 4: Pairwise-based means hypothesis testing p -values for trapezoidal *vs* other LU 's data in 30 simulations of samples of size 100 with the 1st procedure

	Tra vs Π	Tra vs LU_{1A}	Tra vs LU_{1B}	Tra vs LU_{2A}	Tra vs LU_{2B}	Tra vs Tri	Tra vs TriS
SAMPLE 1	0.855	0.452	0.257	0.622	0.140	0.006	0.009
SAMPLE 2	0.932	0.569	0.631	0.870	0.530	0.009	0.015
SAMPLE 3	0.862	0.408	0.412	0.753	0.242	0.000	0.007
SAMPLE 4	0.882	0.498	0.430	0.756	0.277	0.009	0.015
SAMPLE 5	0.938	0.581	0.802	0.915	0.717	0.002	0.003
SAMPLE 6	0.915	0.638	0.511	0.776	0.385	0.003	0.003
SAMPLE 7	0.865	0.263	0.688	0.852	0.513	0.003	0.002
SAMPLE 8	0.932	0.547	0.659	0.862	0.514	0.013	0.032
SAMPLE 9	0.937	0.624	0.690	0.859	0.600	0.021	0.015
SAMPLE 10	0.873	0.500	0.541	0.807	0.432	0.001	0.001
SAMPLE 11	0.868	0.480	0.401	0.732	0.288	0.000	0.001
SAMPLE 12	0.957	0.723	0.862	0.948	0.807	0.040	0.048
SAMPLE 13	0.851	0.358	0.536	0.794	0.398	0.005	0.001
SAMPLE 14	0.932	0.618	0.647	0.855	0.516	0.002	0.006
SAMPLE 15	0.913	0.560	0.720	0.912	0.601	0.001	0.000
SAMPLE 16	0.886	0.494	0.426	0.746	0.265	0.012	0.034
SAMPLE 17	0.937	0.688	0.547	0.822	0.447	0.003	0.011
SAMPLE 18	0.897	0.472	0.681	0.858	0.551	0.020	0.017
SAMPLE 19	0.884	0.427	0.567	0.837	0.427	0.001	0.004
SAMPLE 20	0.887	0.572	0.385	0.743	0.274	0.012	0.016
SAMPLE 21	0.899	0.594	0.507	0.786	0.382	0.004	0.006
SAMPLE 22	0.871	0.373	0.490	0.787	0.339	0.006	0.016
SAMPLE 23	0.893	0.402	0.613	0.826	0.470	0.026	0.033
SAMPLE 24	0.924	0.590	0.592	0.852	0.466	0.006	0.005
SAMPLE 25	0.889	0.364	0.464	0.774	0.314	0.009	0.029
SAMPLE 26	0.874	0.441	0.436	0.767	0.333	0.017	0.020
SAMPLE 27	0.898	0.548	0.551	0.823	0.403	0.004	0.013
SAMPLE 28	0.887	0.468	0.584	0.809	0.426	0.005	0.005
SAMPLE 29	0.919	0.518	0.506	0.809	0.382	0.014	0.030
SAMPLE 30	0.850	0.356	0.320	0.662	0.192	0.006	0.005

very small in testing the equality of means of trapezoidal data *vs* triangular ones. Consequently, we can conclude that there are no evidences against the null hypothesis of the equality of means in the first five cases (actually, those preserving not only the support but also the core), but there are strong evidences against such a hypothesis when fuzzy data are triangular (notice that although the support is shared, the core is reduced to a singleton). Among the first five comparisons, evidences against the equality of means are weaker for the Π and LU_{2A} (which seems to be expected by looking at Figure 1).

According to the SECOND SIMULATION PROCEDURE, data have been generated from random fuzzy numbers with a bounded reference set and abstracting and mimicking what we have observed in conducting different (computerized and paper-and-pencil) questionnaires based on the fuzzy rating scale by Hesketh *et al.* [29] (see, for instance, Colubi *et al.* [11], González-Rodríguez *et al.* [22], and Lubiano *et al.* [34]) and being similar to the ones

followed by De la Rosa de Saa *et al.* [13]. More concretely, fuzzy data have been generated so that

- 5% of the data have been obtained by first considering a simulation from a simple random sample of size 4 from a beta $\beta(1, 1)$ distribution, later scaling values in the sample to (say) interval $[0, 10]$ and considering the ordered sample, and finally computing the values of the x_i .
- 35% of the data have been obtained considering a simulation of four random variables $X_i = 10 \cdot Y_i$ as follows:

$$\begin{aligned} Y_1 &\sim \beta(1, 1), \\ Y_2 &\sim \text{Uniform}[0, \min\{1/10, Y_1, 1 - Y_1\}], \\ Y_3 &\sim \text{Uniform}[0, \min\{1/5, Y_1 - Y_2\}], \\ Y_4 &\sim \text{Uniform}[0, \min\{1/5, 1 - Y_1 - Y_2\}]; \end{aligned}$$

- 60% of the data have been obtained considering a simulation of four random variables $X_i = 10 \cdot Y_i$ as follows:

$$\begin{aligned} Y_1 &\sim \beta(1, 1), \\ Y_2 &\sim \begin{cases} \text{Exp}(200) & \text{if } Y_1 \in [0.25, 0.75] \\ \text{Exp}(100 + 4Y_1) & \text{if } Y_1 < 0.25 \\ \text{Exp}(500 - 4Y_1) & \text{otherwise} \end{cases} \\ Y_3 &\sim \begin{cases} \gamma(4, 100) & \text{if } Y_1 - Y_2 \geq 0.25 \\ \gamma(4, 100 + 4Y_1) & \text{otherwise} \end{cases} \\ Y_4 &\sim \begin{cases} \gamma(4, 100) & \text{if } 1 - Y_1 - Y_2 \leq 0.75 \\ \gamma(4, 500 - 4Y_1) & \text{otherwise.} \end{cases} \end{aligned}$$

When this SECOND PROCEDURE is considered, results have been gathered in Table 5.

The p -values in testing the equality of means of trapezoidal data *vs* any of the other LU -type data are all high in this table. Consequently, we can conclude that there are no evidences against the null hypothesis of the equality of means. Among the seven comparisons, evidences against the equality of means are very weak for the Π and LU_{2A} cases and slightly less weak for the triangular one.

Remark 3.2. Studies similar to those in Tables 4 and 5 have been also developed for other sample sizes as well as for Bertoluzza *et al.*'s L^2 metric [6]

$$D(\tilde{U}, \tilde{V}) = \sqrt{\int_{[0,1]} \left[\int_{[0,1]} [\tilde{U}_\alpha^{[t]} - \tilde{V}_\alpha^{[t]}]^2 dt \right] d\alpha},$$

Table 5: Pairwise-based means hypothesis testing p -values for trapezoidal *vs* other LU 's data in 30 simulations of samples of size 100 with the 2nd procedure

	Tra <i>vs</i> II	Tra <i>vs</i> LU_{1A}	Tra <i>vs</i> LU_{1B}	Tra <i>vs</i> LU_{2A}	Tra <i>vs</i> LU_{2B}	Tra <i>vs</i> Tri	Tra <i>vs</i> TriS
SAMPLE 1	0.946	0.788	0.823	0.904	0.759	0.583	0.557
SAMPLE 2	0.957	0.791	0.831	0.926	0.770	0.654	0.661
SAMPLE 3	0.955	0.810	0.832	0.931	0.748	0.594	0.589
SAMPLE 4	0.959	0.838	0.844	0.935	0.779	0.620	0.600
SAMPLE 5	0.926	0.752	0.787	0.878	0.692	0.613	0.615
SAMPLE 6	0.966	0.835	0.862	0.944	0.799	0.610	0.601
SAMPLE 7	0.954	0.804	0.809	0.904	0.766	0.721	0.696
SAMPLE 8	0.955	0.834	0.812	0.908	0.761	0.643	0.634
SAMPLE 9	0.960	0.825	0.809	0.911	0.741	0.626	0.632
SAMPLE 10	0.950	0.816	0.805	0.912	0.754	0.616	0.640
SAMPLE 11	0.963	0.836	0.824	0.927	0.776	0.568	0.549
SAMPLE 12	0.929	0.705	0.760	0.878	0.719	0.551	0.518
SAMPLE 13	0.949	0.771	0.779	0.896	0.717	0.619	0.627
SAMPLE 14	0.957	0.803	0.836	0.933	0.769	0.597	0.598
SAMPLE 15	0.968	0.836	0.866	0.947	0.790	0.591	0.616
SAMPLE 16	0.972	0.851	0.867	0.939	0.797	0.658	0.661
SAMPLE 17	0.964	0.832	0.847	0.926	0.802	0.617	0.608
SAMPLE 18	0.955	0.793	0.752	0.900	0.703	0.595	0.626
SAMPLE 19	0.947	0.766	0.816	0.913	0.749	0.642	0.651
SAMPLE 20	0.947	0.786	0.785	0.897	0.768	0.632	0.655
SAMPLE 21	0.938	0.758	0.792	0.902	0.714	0.660	0.633
SAMPLE 22	0.933	0.746	0.736	0.869	0.644	0.658	0.606
SAMPLE 23	0.949	0.788	0.816	0.915	0.762	0.601	0.616
SAMPLE 24	0.961	0.787	0.803	0.899	0.746	0.627	0.639
SAMPLE 25	0.952	0.802	0.819	0.928	0.769	0.690	0.673
SAMPLE 26	0.932	0.746	0.770	0.898	0.718	0.639	0.636
SAMPLE 27	0.950	0.789	0.829	0.925	0.797	0.661	0.647
SAMPLE 28	0.936	0.748	0.803	0.890	0.729	0.577	0.579
SAMPLE 29	0.963	0.809	0.840	0.945	0.792	0.584	0.598
SAMPLE 30	0.952	0.785	0.795	0.904	0.703	0.653	0.647

where $\tilde{U}_\alpha^{[t]} = t \cdot \sup \tilde{U}_\alpha + (1 - t) \cdot \inf \tilde{U}_\alpha$. Conclusions do not vary too much. In connection with the change of sample size, one can comment that p -values are mostly higher for smaller sample sizes, for both simulation procedures. In connection with the choice of the metric, the influence is not that big but p -values are generally slightly higher for Bertoluzza *et al.*'s metric to the extent that, in testing *vs* triangular data, 'differences' which are not significant at the significance level .05 with Bertoluzza *et al.*'s metric for some samples from the 1st procedure are associated with significant ones at the same level with ρ_2 .

4. Additional discussion of the influence of the shape of fuzzy data on their mean values through a real-life example

In this section we examine the case study in Gil *et al.* [20] and Lubiano *et al.* [34] regarding some bootstrapped test about means of the one-, two- and

k -samples p -values. Results from this real-life example will confirm, from a slightly different perspective, the conclusions in Section 3, in accordance with which the shape of the fuzzy data does not affect very much the statistical conclusions about means.

The real-life example concerns the questionnaire TIMSS-PIRLS 2011, that was designed to be conducted on fourth grade students and questions referring to their attitude w.r.t. reading, mathematics and science. Most of the involved items had to be answered according to the 4-point Likert scale given by A1 = DISAGREE A LOT, A2 = DISAGREE A LITTLE, A3 = AGREE A LITTLE and A4 = AGREE A LOT.

The problem faced in [20, 34] is that of statistically analyzing whether adapting these questions with the fuzzy rating scale approach would yield somewhat different statistical conclusions. For this reason, the questionnaire form has been adapted to allow a double-type response, namely, the original Likert and a fuzzy rating scale-based one (see Figure 4 for Question *M.2*, “My teacher is easy to understand”).

Mathematics in school

Mathematics

How much do you agree with these statements about learning mathematics?

M.2 . My teacher is easy to understand

1. Disagree a lot — Disagree a little — Agree a little — Agree a lot

2. 0 1 2 3 4 5 6 7 8 9 10

Figure 4: Example of the double-response form to a question

The questionnaire has been conducted in 2014 on fourth grade students from Colegio San Ignacio, Oviedo-Asturias, Spain (see <http://bellman.ciencias.uniovi.es/SMIRE/FuzzyRatingScaleQuestionnaire-SanIgnacio.html> for the full paper-and-pencil and computerized versions). In this way, each of the questions is assumed to be filled in accordance with both the 4-point Likert and the

fuzzy rating scales with reference interval $[0, 10]$ (Spain uses a 10-point grading scale for elementary and high schools, so it could mean a proper simile for students).

The fuzzy rating scale is assumed to be based on the free trapezoidal assessment of responses. Some students have used the computer-administered format, whereas the others have filled the paper-and-pencil one. The training of the 9-year-old children has taken up to fifteen minutes, and there were not incidences to be noticed in conducting the questionnaire. The complete questionnaire and dataset can be found in <http://bellman.ciencias.uniovi.es/SMIRE/FuzzyRatingScaleQuestionnaire-SanIgnacio.html>.

For each of the three bootstrapped testing methods in Section 2 (namely, the one-, the two- and the k -sample), the p -values associated with the sample(s) of trapezoidal data are to be compared with those for the sample(s) of the corresponding LU -type fuzzy data with $LU \in \{\Pi, LU_{1A}, LU_{1B}, LU_{2A}, LU_{2B}, \text{Tri}, \text{TriS}\}$. For a more detailed illustration, two situations have been considered for each method: one leading to small p -values and another one leading to moderate/high p -values.

In connection with the one-sample test we have the following:

Example 4.1. Consider the fuzzy rating scale datasets associated with the responses to item $M.2$ (“My teacher is easy to understand”) for the 45 students who have responded to the question by using the computerized form.

In testing whether the mean of the LU response to $M.2$ equals the ‘ LU -ized’ version of the original fuzzy number which appears in the computerized form as the default answer, $LU(2.5, 3.75, 6.25, 7.5) = LU\langle 5, 1.25, 1.25, 1.25 \rangle$, we get the p -values in Table 6.

Table 6: One-sample hypothesis testing p -values about the equality of the mean response to item $M.2$ in the computerized form and the hypothetical response $LU(2.5, 3.75, 6.25, 7.5)$ for different LU -type of data

Tra	Π	LU_{1A}	LU_{1B}	LU_{2A}	LU_{2B}	Tri	TriS
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

On the other hand, in testing whether the mean of the LU response to $M.2$ equals $LU(5, 6.5, 7.5, 9) = LU\langle 7, 0.5, 1.5, 1.5 \rangle$ we get the p -values in Table 7.

In connection with the two-sample test we have the following:

Table 7: One-sample hypothesis testing p -values about the equality of the mean response to item $M.2$ in the computerized form and the hypothetical response $LU(5, 6.5, 7.5, 9)$ for different LU -type of data

Tra	Π	LU_{1A}	LU_{1B}	LU_{2A}	LU_{2B}	Tri	TriS
0.107	0.141	0.268	0.095	0.134	0.101	0.093	0.085

Example 4.2. Consider the fuzzy rating scale dataset associated with the responses to item $M.2$ (“My teacher is easy to understand”) for the 45 students who have responded to the item by using the computerized form and the dataset for the 23 who have responded to the item by using the paper-and-pencil form.

In testing whether the means of the ‘ LU -ized’ responses to $M.2$ for both forms coincide we get the p -values in Table 8.

Table 8: Two-sample hypothesis testing p -values about the equality of the mean responses to item $M.2$ of students using the computerized *vs* the paper-and-pencil form for different LU -type of data

Tra	Π	LU_{1A}	LU_{1B}	LU_{2A}	LU_{2B}	Tri	TriS
0.015	0.012	0.005	0.012	0.013	0.010	0.009	0.013

On the other hand, consider the fuzzy rating scale datasets associated with the responses to item $M.2$ for the 37 girls and the 31 boys who have responded to this item. In testing whether the means of the ‘ LU -ized’ responses to $M.2$ are irrespective of sex, we get the p -values in Table 9.

Table 9: Two-sample hypothesis testing p -values about the equality of the mean response to item $M.2$ of girls *vs* boys for different LU -type of data

Tra	Π	LU_{1A}	LU_{1B}	LU_{2A}	LU_{2B}	Tri	TriS
0.589	0.545	0.571	0.534	0.563	0.587	0.536	0.572

In connection with the one-way ANOVA test we have the following:

Example 4.3. In the first ANOVA study the considered factor has been ‘mark taken by the students in the last examination of math’ (an additional question included in the conducted student’s questionnaire). This factor has been assumed to act at 4 levels, associated with

$$\mathbf{G1} = [0, 6], \quad \mathbf{G2} = (6, 8], \quad \mathbf{G3} = (8, 9], \quad \mathbf{G4} = (9, 10],$$

the choice of these levels having been based on attempting to get a rather balanced distribution of students among groups. The influence of the factor is to be analyzed on the response to item $M.2$ of the 62 students responding simultaneously to $M.2$ and the additional question leading to assign them with one of the levels of the factor. We get the p -values in Table 10.

Table 10: ANOVA hypothesis testing p -values for the responses to item $M.2$ of students depending on the mark taken in the last examination of math

Tra	Π	LU_{1A}	LU_{1B}	LU_{2A}	LU_{2B}	Tri	TriS
0.011	0.013	0.022	0.011	0.024	0.019	0.027	0.030

In the second ANOVA study the considered factor has been ‘mark taken by the students in the last examination of reading’ (also an additional question included in the conducted student’s questionnaire). This factor has been assumed to act at the above described 4 levels $G1$, $G2$, $G3$, and $G4$. The influence of the factor is to be analyzed on the response to item $R.3$ of the 62 students responding simultaneously to $R.3$ (“Reading is harder for me than any other subject”) and the additional question leading to assign them with one of the levels of the factor. We get the p -values in Table 11.

Table 11: ANOVA hypothesis testing p -values for the responses to item $R.3$ of students depending on the mark taken in the last examination of reading

Tra	Π	LU_{1A}	LU_{1B}	LU_{2A}	LU_{2B}	Tri	TriS
0.471	0.455	0.477	0.459	0.474	0.441	0.470	0.507

As a common conclusion from the last three examples, one can empirically assert that the results for both low and not that low p -values seem not to be very much affected by the shape of the fuzzy responses. This supports and reinforces the conclusions in Section 3. If ρ_2 is replaced by Bertoluzza *et al.*’s metric, conclusions scarcely change.

5. Concluding remarks

The main implication from the analysis in this paper is that the use of trapezoidal data, especially in developments involving the fuzzy rating scale, can be appropriately justified/allowed when we are interested in the statistical conclusions about means. It should be pointed out that the use of

triangular data can be more or less influential on the conclusions when cores are far from singletons.

A closely related open problem is that of analyzing the influence of the choice of the shape of fuzzy data associated with random experiments on other location/central tendency or dispersion/scale measures of the distribution of random fuzzy numbers, like for instance, the robust measures by Sinova *et al.* [39, 40, 41] and De la Rosa de Súa *et al.* [14].

Acknowledgements

The research in this paper has been partially supported by the Principality of Asturias/FEDER Grant GRUPIN14-101. Its financial support is gratefully acknowledged. The authors are also grateful to their colleagues Sara de la Rosa de Súa and Beatriz Sinova for their helpful comments in connection with this paper.

References

- [1] S. Abbasbandy, B. Asady, The nearest trapezoidal fuzzy number to a fuzzy quantity, *Appl. Math. Comput.* 156 (2004) 381–386.
- [2] S. Baglio, L. Fortuna, S. Graziani, G. Muscato, Membership function shape and the dynamical behaviour of a fuzzy system, in: *Proc. EUFIT'93, Aachen, 1993*, pp. 645–650.
- [3] A.I. Ban, L. Coroianu, Nearest interval, triangular and trapezoidal approximation of a fuzzy number preserving ambiguity, *Int. J. Approx. Reason.* 53 (2012) 805–836.
- [4] A.I. Ban, L. Coroianu, Existence, uniqueness and continuity of trapezoidal approximations of fuzzy numbers under a general condition, *Fuzzy Sets Syst.* 257 (2014) 3–22.
- [5] R. Belohlavek, Do exact shapes of fuzzy sets matter? *Int. J. Gen. Syst.* 36(5) (2007) 527–555.
- [6] C. Bertoluzza, N. Corral, A. Salas, On a new class of distances between fuzzy numbers, *Mathware Soft Comput.* 2 (1995) 71–84.
- [7] A. Blanco-Fernández, M.R. Casals, A. Colubi, N. Corral, M. García-Bárzana, M.A. Gil, G. González-Rodríguez, M.T. López, M.A. Lubiano, M. Montenegro, A.B. Ramos-Guajardo, S. De la Rosa de Súa, B. Sinova, A distance-based statistical analysis of fuzzy number-valued data. *Int. J. Approx. Reason.* 55 (2014) 1487–1501.
- [8] A. Blanco-Fernández, M.R. Casals, A. Colubi, N. Corral, M. García-Bárzana, M.A. Gil, G. González-Rodríguez, M.T. López, M.A. Lubiano, M. Montenegro, A.B. Ramos-Guajardo, S. De la Rosa de Súa, B. Sinova, Rejoinder on “A distance-based statistical analysis of fuzzy number-valued data.” *Int. J. Approx. Reason.* 55 (2014) 1601–1605.

- [9] A. Colubi, Statistical inference about the means of fuzzy random variables: Applications to the analysis of fuzzy- and real-valued data, *Fuzzy Sets Syst.* 160 (2009) 344–356.
- [10] A. Colubi, J.S. Domínguez-Menchero, M. López-Díaz, D.A. Ralescu, On the formalization of fuzzy random variables, *Inform. Sciences* 133 (2001) 3–6.
- [11] A. Colubi, G. González-Rodríguez, Ma.A. Gil, W. Trutschnig, Nonparametric criteria for supervised classification of fuzzy data, *Int. J. Approx. Reason.* 52 (2011) 1272–1282.
- [12] L. Coroianu, L. Stefanini, General approximation of fuzzy numbers by F -transform, *Fuzzy Sets Syst.* 288 (2016) 46–74.
- [13] S. de la Rosa de Sáa, M.A. Gil, G. González-Rodríguez, M.T. López, M.A. Lubiano, Fuzzy rating scale-based questionnaires and their statistical analysis, *IEEE Trans. Fuzzy Syst.* 23(1) (2015) 111–126.
- [14] S. de la Rosa de Sáa, M.A. Lubiano, B. Sinova, P. Filzmoser, Robust scale estimators for fuzzy data, *Adv. Data Anal. Classif.* (in press, doi:10.1007/s11634-015-0210-1).
- [15] T. Denoeux, Maximum likelihood estimation from fuzzy data using the EM algorithm, *Fuzzy Sets Syst.* 183 (2011)72–91.
- [16] P. Diamond, P. Kloeden, Metric spaces of fuzzy sets, *Fuzzy Sets Syst.* 35 (1990) 241–249.
- [17] J. Dombi, Z. Gera, The approximation of piecewise linear membership functions and Lukasiewicz operators, *Fuzzy Sets Syst.* 154 (2005) 275–286.
- [18] A. Ebrahimnejad, J.L. Verdegay, A novel approach for sensitivity analysis in linear programs with trapezoidal fuzzy numbers, *J. Intell. Fuzzy Syst.* 27 (2014) 173–185.
- [19] C. García-Aguado, J.L. Verdegay, On the sensitivity of membership functions for fuzzy linear programming problems, *Fuzzy Sets Syst.* 56 (1993) 47–49.
- [20] M.A. Gil, M.A. Lubiano, S. de la Rosa de Sáa and B. Sinova, Analyzing data from a fuzzy rating scale-based questionnaire. A case study, *Psicothema* 27 (2015) 182–191.
- [21] M.A. Gil, M. Montenegro, G. González-Rodríguez, A. Colubi, M.R. Casals, Bootstrap approach to the multi-sample test of means with imprecise data, *Comput. Stat. Data Anal.* 51 (2006) 148–162.
- [22] G. González-Rodríguez, A. Colubi, M.A. Gil, Fuzzy data treated as functional data: A one-way ANOVA test approach, *Comput. Stat. Data An.* 56 (2012) 943–955.
- [23] G. González-Rodríguez, M. Montenegro, A. Colubi, M.A. Gil, M.A. (2006). Bootstrap techniques and fuzzy random variables: Synergy in hypothesis testing with fuzzy data, *Fuzzy Sets Syst.* 157 (2006) 2608–2613.
- [24] P. Grzegorzewski, Trapezoidal approximations of fuzzy numbers preserving the expected interval – Algorithms and properties, *Fuzzy Sets Syst.* 159 (2008) 1354–1364.
- [25] P. Grzegorzewski, E. Mrówka, Trapezoidal approximations of fuzzy numbers, *Fuzzy Sets Syst.* 153 (2005) 115–135.
- [26] P. Grzegorzewski, K. Pasternak-Winiarska, Natural trapezoidal approximations of fuzzy numbers, *Fuzzy Sets Syst.* 250 (2014) 90–109.

- [27] P. Grzegorzewski, L. Stefanini, Fuzzy numbers and their applications, *Fuzzy Sets Syst.* 257 (2014) 1–2.
- [28] T. Hesketh, B. Hesketh, Computerized fuzzy ratings: the concept of a fuzzy class. *Behav. Res. Meth. Ins. C.* 26 (1994) 272–277.
- [29] T. Hesketh, R. Pryor, B. Hesketh, An application of a computerized fuzzy graphic rating scale to the psychological measurement of individual differences, *Int. J. Man-Mach. Stud.* 29 (1988) 21–35.
- [30] P.D. Koprinkova, Membership functions shape and its influence on the stability of fuzzy control systems, *Cybernet. Syst.* 31(4) (2000) 353371.
- [31] P.D. Koprinkova-Hristova, V. Penev, Dynamical behavior of fuzzy logic based velocity control autopilot with respect to changes in linguistic variables membership functions shape, *Inf. & Secur.* 3 (1999) 108–115.
- [32] R. Körner, An asymptotic α -test for the expectation of random fuzzy variables, *J. Stat. Plan. Infer.* 83 (2000) 331–346.
- [33] R. Kruse, K.D. Meyer, *Statistics with Vague Data*. Springer-Theory and Decision Library B (formerly Reidel Publ. Co.), Dordrecht, 1987.
- [34] M.A. Lubiano, M. Montenegro, B. Sinova, S. de la Rosa de Sáa, M.A. Gil, Hypothesis testing for means in connection with fuzzy rating scale-based data: algorithms and applications, *Eur. J. Oper. Res.* 251 (2016) 918–929.
- [35] M. Montenegro, M.R. Casals, M.A. Lubiano, M.A. Gil, Two-sample hypothesis tests of means of a fuzzy random variable, *Inform. Sciences* 133 (2001) 89–100.
- [36] M. Montenegro, A. Colubi, M.R. Casals, M.A. Gil, Asymptotic and Bootstrap techniques for testing the expected value of a fuzzy random variable, *Metrika* 59 (2004) 31–49.
- [37] M.L. Puri, D.A. Ralescu, Fuzzy random variables, *J. Math. Anal. Appl.* 114 (1986) 409–422.
- [38] D.Z. Šaletić, D.M. Velašević, N.E. Mastorakis, The influence of shapes of fuzzy sets membership functions on fuzzy system performances, *Int. J. Comput. Res.* 11(2) (2002) 111–125.
- [39] B. Sinova, S. de la Rosa de Sáa, M.A. Gil, A generalized L^1 -type metric between fuzzy numbers for an approach to central tendency of fuzzy data. *Inform. Sciences* 242 (2013) 22–34.
- [40] B. Sinova, M.A. Gil, A. Colubi, S. Van Aelst, The median of a random fuzzy number. The 1-norm distance approach. *Fuzzy Sets Syst.* 200 (2012) 99–115.
- [41] B. Sinova, M.A. Gil, S. Van Aelst, M-estimates of location for the robust central tendency of fuzzy data, *IEEE Trans. Fuzzy Syst.* 24 (2016) 945–956.
- [42] L. Sorini, L. Stefanini, Some parametric forms for LR fuzzy numbers and LR fuzzy arithmetic, in: *Proc. ISDA '09, Pisa, 2009*, pp. 312–317.
- [43] L. Stefanini, B. Bede, Generalized fuzzy differentiability with LU-parametric representation, *Fuzzy Sets Syst.* 257 (2014) 184–203.
- [44] L. Stefanini, L. Sorini, Fuzzy arithmetic with arithmetic LR fuzzy numbers, in: *Proc. IFSA/EUSFLAT, Lisbon, 2009*, pp. 600–605.

- [45] L. Stefanini, L. Sorini, M.L. Guerra, Parametric representation of fuzzy numbers and applications to fuzzy calculus. *Fuzzy Sets Syst.* 157 (2006) 2423–2455.
- [46] W. Trutschnig, M.A. Lubiano, SAFD: Statistical Analysis of Fuzzy Data. R package version 1.0-1 (<http://cran.r-project.org/web/packages/SAFD/index.html>, version 2015-12-10).
- [47] J.L. Verdegay, Progress on Fuzzy Mathematical programming: a personal perspective, *Fuzzy Sets Syst.* 281 (2015) 219–226.
- [48] R.A. Vitale, L_p metrics for compact, convex sets, *J. Approx. Theor.* 45 (1985) 280–287.
- [49] C.-T. Yeh, Trapezoidal and triangular approximations preserving the expected interval, *Fuzzy Sets Syst.* 159 (2008) 1345–1353.
- [50] C.-T. Yeh, Weighted semi-trapezoidal approximations of fuzzy numbers, *Fuzzy Sets Syst.* 165 (2011) 61–80.
- [51] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, Part 1, *Inform. Sciences* 8 (1975) 199–249; Part 2, *Inform. Sciences* 8 (1975) 301–353; Part 3, *Inform. Sciences* 9 (1975) 43–80.